LECTURE 3. HYPERBOLIC GEOMETRY AND KLEINIAN GROUPS

3.1 The hyperbolic plane: half-plane and disc models, isometries

Around 300BC Euclid of Alexandria wrote a thirteen volume treatise entitled *The Elements*, in which he developed geometry and number theory from a set of *axioms*. His five axioms for geometry in the plane were:

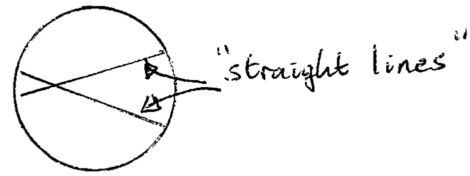
1. A straight line may be drawn from any point to any other point.
2. A finite straight line may be extended.
3. A circle may be drawn with any given centre and radius.
4. All right angles are equal.
5. If a straight line intersects two other straight lines and the sum of the interior angles on one side is less than two right angles, then the two straight lines, if extended indefinitely, meet on the side on which the sum of the angles is less than two right angles.

The fifth postulate is equivalent to:

5′. Given any straight line and point not on it, there exists a unique straight line through the point, not meeting the given line.

There were many attempts in the following two thousand years to show that the fifth axiom can be deduced from the first four. It appears to be Gauss who was the first to realise that there existed a geometry satisfying axioms 1 to 4 but not 5. He called this a *non-Euclidean geometry*, but though he investigated its properties for ten years in the early 19th century, he did not publish any of his results. It was Lobachevsky (1829) and Bolyai (1832) who first published the discovery of what we now call *hyperbolic geometry*, which has in place of Euclid's axiom that 'there exists a unique parallel', the new axiom that 'there exist infinitely many parallels' to a given line, through a given point not on it.

Beltrami (1868) introduced a Euclidean disc model of the hyperbolic plane



which he used to prove formally that Euclid's fifth axiom is independent of the first four. Klein (1871) gave an interpretation of this model in terms of projective geometry. Beltrami had also introduced conformal disc and upper half-plane models in 1868,

and Poincaré (1882) identified the congruences of the hyperbolic plane with the group $PSL(2, \mathbf{R})$ of the upper half-plane, the key to a host of subsequent developments in mathematics and physics in the subsequent century (from relativity to string theory).

The upper half-plane model

$$\mathcal{H}^2 = \mathcal{H}_+ = \{x + iy : x \in \mathbf{R}, y \in \mathbf{R}^{>0}\} \subset \mathbf{C}$$

Define an infinitesimal metric on $\mathcal{H}^+$ by

$$ds = \frac{1}{y}((dx)^2 + (dy)^2)^{1/2}$$

in other word the 'length' of a path γ in $\mathcal{H}_+$ is defined to be the integral of this quantity ds along γ.

Lemma 3.1 *ds is invariant under $PSL(2,\mathbf{R})$.*
Proof

$$z \to \frac{az+b}{cz+d} = \frac{a}{c} + \frac{r}{cz+d}$$

where $r = b - ad/c$, so it suffices to check invariance under the following three types of transformation: (i) $z \to z + \lambda \quad (\lambda \in \mathbf{R})$; (ii) $z \to \lambda z \quad (\lambda \in \mathbf{R}^{>0})$; (iii) $z \to -1/z$. This is an easy exercise. QED

Definition A path γ is called a *geodesic* from P to Q in $\mathcal{H}_+$ if it is a path of shortest length from P to Q. A proof of the following elementary proposition can be found in any textbook on hyperbolic geometry.

Proposition 3.2 *There is a unique geodesic between any two distinct points P and Q in $\mathcal{H}_+$. It is the segment between P and Q of the unique (Euclidean) semicircle through P and Q which meets $\hat{\mathbf{R}} = \mathbf{R} \cup \infty$ orthogonally. The (hyperbolic) distance from P to Q is $|\ln(P,Q;A,B)|$ where A and B are the points where the semicircle meets $\hat{\mathbf{R}}$.*

Any isometry of the hyperbolic plane $\mathcal{H}^2$ must send geodesics to geodesics, and so in the upper half plane model it must send semicircles orthogonal to $\hat{\mathbf{R}}$ to semicircles orthogonal to $\hat{\mathbf{R}}$. It can be shown that such transformations must have *either* the form

$$z \to \frac{az+b}{cz+d} \quad a,b,c,d \in \mathbf{R}, ad - bc > 0$$

or the form

$$z \to \frac{a\bar{z}+b}{c\bar{z}+d} \quad a,b,c,d \in \mathbf{R}, ad - bc < 0$$

An example of the second type is $z \to -\bar{z}$ (reflection in the imaginary axis).

The disc model

An alternative model to the upper half plane is given by regarding $\mathcal{H}^2$ as the points of the unit disc $D^2 = \{z : |z| < 1\} \subset \mathbf{C}$ and taking as the infinitesimal metric on this disc

$$ds = \frac{((dx)^2 + (dy)^2)^{1/2}}{1 - r^2} \qquad (r^2 = x^2 + y^2)$$

in which case the geodesics are arcs of circles meeting the unit circle (the boundary of D^2) orthogonally. To pass from the upper half plane model to the disc model to the disc model we simply conjugate by any element of $PSL(2, \mathbf{C})$ which sends $\mathcal{H}_+$ to D^2, for example the map

$$z \to \frac{iz + 1}{z + i}$$

which sends $-1, 0, 1$ to $-1, -i, 1$ respectively and hence sends $\hat{\mathbf{R}}$ to the unit circle (and moreover sends the upper half plane to the interior of this circle since it sends i to 0).

The *orientation-preserving isometries* of the hyperbolic plane are the elements of $PSL(2, \mathbf{R})$, acting as fractional linear maps

$$z \to \frac{az + b}{cz + d}$$

in the upper half-plane model. We use their fixed points to classify them into types.

Definition A non-identity element $\alpha \in PSL(2, \mathbf{R})$ is said to be
elliptic if it has a fixed point in $\mathcal{H}_+$;
parabolic if it has precisely one fixed point on $\hat{\mathbf{R}}$;
hyperbolic if has two fixed points on $\hat{\mathbf{R}}$.

If we normalise our matrix representing $\alpha \in PSL(2, \mathbf{R})$ so that $ad - bc = 1$, we can distinguish the three types by the *trace*, $a + d$ of α as follows. The fixed points of α are the solutions of the equation

$$cz^2 + (d + a)z - b = 0$$

This has a complex conjugate pair of roots $\Leftrightarrow (d + a)^2 + 4bc < 0 \Leftrightarrow (d + a)^2 - 4 < 0 \Leftrightarrow |tr(\alpha)| < 2$, it has one (repeated) real root $\Leftrightarrow |tr(\alpha)| = 2$, and it has two (distinct) real roots $\Leftrightarrow |tr(\alpha)| > 2$. Thus

Lemma 3.3 *α is elliptic/parabolic/hyperbolic* $\Leftrightarrow |tr(\alpha)| < 2, = 2, > 2$

(Note that we must normalise α to determinant 1 before we compute the trace.)

The trace of a matrix is a conjugacy invariant, and hence so is the *type* of an isometry of the hyerbolic plane (this is also obvious from the definition of *type* in terms of fixed points). In calculations and proofs it can often be useful to conjugate an isometry to a standard form. The following can easily be verified:

Lemma 3.4 *If $\alpha \in PSL(2, \mathbf{R})$ is parabolic then α is conjugate (in $PSL(2, \mathbf{R})$) to $z \to z+1$ or to $z \to z - 1$.*

Lemma 3.5 *In the disc model, the elliptic elements fixing the origin* 0 *are the Euclidean rotations of the disc.*

Lemma 3.6 *If* $\alpha \in PSL(2, \mathbf{R})$ *is hyperbolic then* α *is conjugate (in* $PSL(2, \mathbf{R})$*) to* $z \to \lambda z$ *for some* $\lambda \in \mathbf{R}^{>0}$.

We may change λ to λ^{-1} by further conjugating our map by $z \to 1/z$ (interchanging 0 and ∞), but since the eigenvalues of α are λ and λ^{-1}, and these are conjugacy invariants of α, the value of λ in Lemma 3.6 is unique up to replacement by λ^{-1}.

3.2 Hyperbolic 3-space and its isometries

Definition $\mathcal{H}^3 = \{(x_1, x_2, x_3) \in \mathbf{R}^3 : x_3 > 0\}$

Just as in the two-dimensional case we may define an infinitesimal metric:

$$ds = \frac{1}{x_3}((dx_1)^2 + (dx_2)^2 + (dx_3)^2)^{1/2}$$

With this metric $\mathcal{H}^3$ becomes the *upper half-space model of hyperbolic* 3*-space.* The geodesics are the semicircles in $\mathcal{H}^3$ orthogonal to the plane $x_3 = 0$.

Now think of the plane $x_3 = 0$ in $\mathbf{R}^3$ as the complex plane $\mathbf{C}$ ($(x_1, x_2, 0) \leftrightarrow x_1 + ix_2$), add the point '$\infty$', and think of $\hat{\mathbf{C}}$ as the *boundary* of $\mathcal{H}^3$. Every fractional linear map

$$\alpha : z \to \frac{az + b}{cz + d} \quad (a, b, c, d \in \mathbf{C}, ad - bc \neq 0)$$

mapping $\hat{\mathbf{C}}$ to $\hat{\mathbf{C}}$, has an extension to an isometry from $\mathcal{H}^3$ to $\mathcal{H}^3$. One way to see this is to break down α into a composition of maps of the form

(*i*)
$$z \to z + \lambda \quad (\lambda \in \mathbf{C})$$

(*ii*)
$$z \to \lambda z \quad (\lambda \in \mathbf{C})$$

(*iii*)
$$z \to -1/z$$

We extend these as follows on $\mathcal{H}^3$ (where z denotes $x_1 + ix_2$):

(*i*)
$$(z, x_3) \to (z + \lambda, x_3)$$

(*ii*)
$$(z, x_3) \to (\lambda z, |\lambda| x_3)$$

(*iii*)
$$(z, x_3) \to \left(\frac{-\bar{z}}{|z|^2 + x_3^2}, \frac{x_3}{|z|^2 + x_3^2}\right)$$

The expressions above come from decomposing the action on $\hat{\mathbf{C}}$ of each of the elements of $PSL(2,\mathbf{C})$ in question into two *inversions* (reflections) in circles in $\hat{\mathbf{C}}$. Each such inversion has a unique extension to $\mathcal{H}^3$ as an inversion in the hemisphere spanned by the circle and composing appropriate pairs of inversions gives us these formulae. It is now an exercise along the lines of Lemma 3.1 to show that $PSL(2,\mathbf{C})$ preserves the metric ds on $\mathcal{H}^3$ and another exercise, along the lines of Lemma 3.2 to show that the geodesics are the arcs of semicircles as claimed. Moreover every isometry of $\mathcal{H}^3$ can be seen to be the extension of a conformal map of $\hat{\mathbf{C}}$ to itself, since it sends hemispheres orthogonal to $\hat{\mathbf{C}}$ to hemispheres orthogonal to $\hat{\mathbf{C}}$, hence circles in $\hat{\mathbf{C}}$ to circles in $\hat{\mathbf{C}}$. Thus all orientation-preserving isometries of $\mathcal{H}^3$ are given by elements of $PSL(2,\mathbf{C})$ acting as above, and all orientation-reversing isometries are extensions of anti-holomorphic Möbius transformations of $\hat{\mathbf{C}}$.

Comments

1. The fact that the orientation-preserving isometry group of $\mathcal{H}^3$ is $PSL(2,\mathbf{C})$ was first observed by Poincaré.
2. The *disc model* for hyperbolic three-space is the interior D of the unit disc in Euclidean three-space $\mathbf{R}^3$, equipped with the metric

$$ds = \frac{((dx_1)^2 + (dx_2)^2 + (dx_3)^2))^{1/2}}{1 - r^2}$$

(where $r^2 = x_1^2 + x_2^2 + x_3^2$). Geodesics are arcs of circles orthogonal to the boundary sphere.
3. One can construct higher dimensional hyperbolic spaces $\mathcal{H}^n$ in the analagous way. In each case the *conformal* transformations of the boundary extend uniquely to give the *isometries* of the interior.

Types of isometries of hyperbolic 3-space

Non-identity elements $\alpha \in PSL(2,\mathbf{C})$ are of four types.

Definition α is said to be

elliptic $\Leftrightarrow$ α fixes some geodesic in $\mathcal{H}^3$ pointwise;

parabolic $\Leftrightarrow$ α has a single fixed point in $\hat{\mathbf{C}}$;

hyperbolic $\Leftrightarrow$ α has two fixed points in $\hat{\mathbf{C}}$, no fixed points in $\mathcal{H}^3$, and every hyperplane in $\mathcal{H}^3$ which contains the geodesic joining the two fixed points in $\hat{\mathbf{C}}$ is invariant (mapped to itself) under α;

loxodromic $\Leftrightarrow$ α has two fixed points in $\hat{\mathbf{C}}$, no fixed points in $\mathcal{H}^3$, and no invariant hyperplane in $\mathcal{H}^3$.

Note The distinction between *hyperbolic* and *loxodromic* is not always made: some authors use either word for an isometry having two fixed points in $\hat{\mathbf{C}}$ and none in $\mathcal{H}^3$.

Lemma 3.7 *α is elliptic/parabolic/hyperbolic/loxodromic*
$\Leftrightarrow (tr(\alpha))^2 \in [0,4) \subset \mathbf{R}^{\geq 0}, = 4, \in \mathbf{R}^{\geq 0} - [0,4), \in \mathbf{C} - \mathbf{R}^{\geq 0}$

Proof

If α has two fixed points in $\hat{\mathbf{C}}$ we may assume (after conjugating α by an appropriate Möbius transformation) they are at 0 and ∞ and that α has the form $z \to \lambda z$ (and $tr(\alpha) = \lambda^{1/2} + \lambda^{-1/2}$).

Case 1: $|\lambda| = 1$, say $\lambda = e^{i\theta}$. Then on $\hat{\mathbf{C}}$ α is a rotation about 0 through an angle θ, and fixes the x_3-axis in $\mathcal{H}^3$ pointwise. As a matrix, normalised to determinant 1,

$$\alpha = \begin{pmatrix} e^{i\theta/2} & 0 \\ 0 & e^{-i\theta/2} \end{pmatrix}$$

and so $(tr(\alpha))^2 = 4\cos^2(\theta/2) \in [0, 4)$.

Case 2: $|\lambda| \neq 1$. then α acts on the x_3-axis in $\mathcal{H}^3$ as multiplication by $|\lambda|$. Writing $\lambda = |\lambda|e^{i\theta}$ we have

$$\alpha = \begin{pmatrix} |\lambda|^{1/2}e^{i\theta/2} & 0 \\ 0 & |\lambda|^{-1/2}e^{-i\theta/2} \end{pmatrix}$$

so $(tr(\alpha))^2 \in \mathbf{C} - [0, 4]$. Now if λ is real (i.e. $\theta = 0$ or π) α is hyperbolic and $(tr(\alpha))^2 \in \mathbf{R}^{\geq 0} - [0, 4]$ and if λ is not real, α is loxodromic and $(tr(\alpha))^2 \in \mathbf{C} - \mathbf{R}^{\geq 0}$.

Finally if α has a single fixed point in $\hat{\mathbf{C}}$ then we can place this fixed point at ∞ (by conjugating α if necessary) in which case α has the form $z \to z + \lambda$ (indeed we may even conjugate it to $z \to z + 1$). Then α is parabolic and $(tr(\alpha))^2 = 4$. QED.

Dynamics of Möbius transformations on $\mathcal{H}^3 \cup \hat{\mathbf{C}}$

$z \to e^{2\pi i\theta}z$ (θ real)

Here the fixed points $0, \infty$ on $\hat{\mathbf{C}}$ are *neutral*. For $z \to e^{i\theta}z$ with θ real, all orbits on $\mathcal{H}^3$ have finite period if θ is a rational multiple of π, and densely fill circles around the x_3 axis if not.

$z \to ke^{2\pi i\theta}z$ ($k > 1, \theta$ real)

Here all orbits in $\mathcal{H}^3$ head away from a repelling fixed point 0 and towards an attracting fixed point ∞, spiralling around the x_3 axis as they go. The nature of the spiralling depends on θ: in particular if $\theta = 0$ or π each orbit remains in a hyperplane.

$z \to z + 1$

In this example the (unique) fixed point ∞ is neutral (multiplier 1) and all orbits on $\mathcal{H}^3$ head towards the fixed point under both forward and backward time. Any parabolic map α will have this behaviour.

3.3 Kleinian groups, ordinary and limit sets, and their properties

Definition A *Kleinian group* is a *discrete* subgroup $G < PSL(2, \mathbf{C})$.

Thus for a subgroup $G < PSL(2, \mathbf{C})$ to be called Kleinian we require that there be no sequence $\{g_n\}$ of distinct elements of G tending to a limit $g \in PSL(2, \mathbf{C})$. Here the topology on $PSL(2, \mathbf{C})$ is that induced by the norm

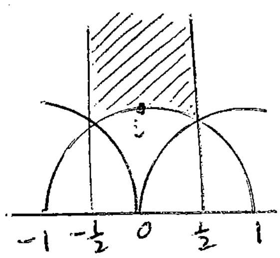
$$|| \begin{pmatrix} a & b \\ c & d \end{pmatrix} || = \sqrt{|a|^2 + |b|^2 + |c|^2 + |d|^2}$$

on $SL(2, \mathbf{C})$ (so that two elements of $PSL(2, \mathbf{C})$ are close together if and only if they are representable by $A_1, A_2 \in SL(2, \mathbf{C})$ with $||A_2 - A_1||$ small).

Note If G is discrete then for any $N > 0$ the number of elements of G having norm $\leq N$ is *finite*, since every infinite sequence with bounded norm has a convergent subsequence. Hence every discrete G is *countable.*

Definition The action of G is *discontinuous* at $z \in \hat{\mathbf{C}}$ if there exists a neighbourhood U of z such that $g(U) \cap U = \emptyset$ for all but finitely many $g \in G$.

Example



$G = PSL(2, \mathbf{Z})$ acts discontinuously on $\hat{\mathbf{C}} - \hat{\mathbf{R}}$. For z in the shaded region above, each $z \neq i, \pm 1/2 + i\sqrt{3}/2$ has a neighbourhood U such that $g(U) \cap U = \emptyset$ for all non-identity $g \in G$, the point $z = i$ has a neighbourhood U such that $g(U) \cap U = \emptyset$ for all $g \in G - \{I, S\}$ where $S : z \to -1/z$, and the point $z = -1/2 + i\sqrt{3}/2$ has a neighbourhood U such that $g(U) \cap U = \emptyset$ for all $g \in G - \{I, ST, (ST)^2\}$ where $ST : z \to -1/(z + 1)$, etc.

Definition The set of all $z \in \hat{\mathbf{C}}$ at which the action of G is discontinuous is called the *regular* (or *ordinary* or *discontinuity*) set $\Omega(G)$.

Comments

1. It follows at once from the definition that $\Omega(G)$ is *open* and *G-invariant.*

2. In the example above observe that the origin 0 is not in $\Omega(G)$, since any U containing 0 has $g(U) \cap U \neq \emptyset$ for all

$$g = \begin{pmatrix} 1 & 0 \\ n & 1 \end{pmatrix}$$

with n sufficiently large. In fact in this example $\Omega(G) = \hat{\mathbf{C}} - \hat{\mathbf{R}}$ (as we shall prove later).

A subgroup $G < PSL(2, \mathbf{C})$ acts on $\mathcal{H}^3$ as well as on its boundary $\hat{\mathbf{C}}$. The following theorem establishes an important relationship between these actions.

Theorem 3.8 *A subgroup $G < PSL(2, \mathbf{C})$ is discrete if and only if it acts discontinuously on $\mathcal{H}^3$*

Proof. If G is not discrete there exists $\{g_n\} \in G$ with limit $g \in PSL(2, \mathbf{C})$. So for all $x \in \mathcal{H}^3$, $g_m^{-1} g_n(x) \to x$ as $m, n \to \infty$. Thus for any $x \in \mathcal{H}^3$ and neighbourhood U of x, for m and n sufficiently large $g_m^{-1} g_n(U) \cap U \neq \emptyset$. Hence G does not act discontinuously at x.

Conversely, if G does not act discontinuously at $x \in \mathcal{H}^3$, then for any neighbourhood U of x there exist a sequence $\{x_n\} \in U$ and (distinct) $g_n \in G$ such that each $g_n(x_n) \in U$. Take U compact. Then by passing to subsequences we may assume the x_n tend to a point y and the $g_n(x_n)$ tend to a point z (with both y and z in U). Now let k be an isometry of $\mathcal{H}^3$ having $k(z) = y$ and let $\{h_n\}$, $\{j_n\}$ be sequences of isometries, both tending to the identity, and having $h_n(y) = x_n$ and $j_n g_n(x_n) = z$ respectively. Consider $f_n = k j_n g_n h_n$. For each n this fixes y (by construction). But the isometries of $\mathcal{H}^3$ fixing a common point of $\mathcal{H}^3$ are a compact group (the Euclidean rotations, in the Poincaré disc model with the common point the origin). Hence the $\{f_n\}$ have a convergent subsequence. Hence so do the $\{g_n\}$, in other words G is not discrete. QED

Limit sets of Kleinian groups

One can define the notion of the *limit set* $\Lambda(G)$ of a Kleinian group G, either in terms of its action on $\mathcal{H}^3$, or in terms of the action on the boundary $\hat{\mathbf{C}}$ of $\mathcal{H}^3$. We shall see later that the two definitions are equivalent.

Definition 1. Let x be any point of $\mathcal{H}^3$. Then set

$$\Lambda(x) = \{w \in \hat{\mathbf{C}} : \exists g_n \in G \quad \text{with} \quad g_n(x) \to w \quad \text{as} \quad n \to \infty\}$$

(where convergence is taken in the Euclidean metric on the Poincaré disc model of $\mathcal{H}^3$). Note that the $\{g_n(x)\}$ cannot have accumulation points in $\mathcal{H}^3$, since G acts discontinuously there. Thus an alternative description of $\Lambda(x)$ is as the accumulation set in $\mathcal{H}^3 \cup \hat{\mathbf{C}}$ of the orbit Gx on $\mathcal{H}^3$. This accumulation set is independent of the initial point $x \in \mathcal{H}^3$, since if we choose another initial point y the hyperbolic distance from $g(x)$ to $g(y)$ is constant for all g and therefore the *Euclidean* distance from $g(x)$ to $g(y)$ tends to zero as $g(x)$ and $g(y)$ approach the boundary $\hat{\mathbf{C}}$ of the Poincaré disc. We *define* $\Lambda(G)$ to be $\Lambda(x)$ for any $x \in \mathcal{H}^3$.

Definition 2. Let z be any point of $\hat{\mathbf{C}}$. Set

$$\Lambda(z) = \{w \in \hat{\mathbf{C}} : \exists g_n \in G \quad \text{with} \quad g_n(z) \to w \quad \text{as} \quad n \to \infty\}$$

(where convergence is taken in the spherical metric on $\hat{\mathbf{C}}$). It can be shown that when G is *non-elementary* (see below for definition) $\Lambda(z)$ is independent of $z \in \hat{\mathbf{C}}$. We define $\Lambda(G)$ to be $\Lambda(z)$ for any $z \in \hat{\mathbf{C}}$.

Comments

1. The restriction that G be 'non-elementary' is included in definition 2 in order to exclude just one class of examples where the limit $\Lambda(z)$ depends on z. Consider $G = \{g^n : n \in \mathbf{Z}\}$, where g is loxodromic, with fixed points z_0 and z_1. The limit set by definition 1 is $\Lambda(G) =$

$\{z_0\} \cup \{z_1\}$, but definition 2 gives $\Lambda(z_0) = z_0$, $\Lambda(z_1) = z_1$ (although $\Lambda(z) = \{z_0\} \cup \{z_1\}$ for any other choice of z).

2. We shall adopt definition 2 until we have proved the equivalence of the two notions (later in this section). Meanwhile we remark that the underlying reason that the definitions are equivalent is that to an observer inside $\mathcal{H}^3$ an orbit of G of $\mathcal{H}^3$ is viewed as accumulating at $\Lambda(G)$ on the 'visual sphere' $\hat{\mathbf{C}}$.

3. A third equivalent definition is that $\Lambda(G)$ consists of the points $z \in \hat{\mathbf{C}}$ where the family $g \in G$ fail to be a normal family (with respect, as always, to the spherical metric). We shall prove this too later in the present section.

4. It follows at once from definition 2 (or indeed from definition 1) that $\Lambda(G)$ is both *closed* and *G-invariant.*

It is clear from the definitions of $\Omega(G)$ and $\Lambda(G)$ that $\Omega(G) \cap \Lambda(G) = \emptyset$, but we shall prove the stronger statement that $\Lambda(G)$ is the *complement* of $\Omega(G)$ in $\hat{\mathbf{C}}$. First we deal with some special cases.

Elementary Kleinian groups

Definition A Kleinian group G is called *elementary* if there exists a finite G orbit on either $\mathcal{H}^3$ or $\hat{\mathbf{C}}$.

All elementary Kleinian groups G belong to the following three classes. For a proof see for example Beardon's book 'Geometry of Discrete Groups' or Ratcliffe's book 'Foundations of Hyperbolic Manifolds.'

(i) G is conjugate to a finite subgroup of $SO(3)$ acting on the Poincaré disc by rigid rotations fixing the origin (for example the symmetry group of a regular solid). In this case $\Lambda(G) = \emptyset$.

(ii) G is conjugate to a discrete group of Euclidean motions of $\mathbf{C}$ (i.e. fixing $\infty \in \hat{\mathbf{C}}$) (for example $G = < z \to z+1, z \to z+i >$). Then $|\Lambda(G)| = 1$.

(iii) G is conjugate to a group all of the elements of which are of the form $z \to kz$ or $z \to k/z$ for $k \in \mathbf{C}$. Then $|\Lambda(G)| = 2$.

It is not hard to see that if G is Kleinian then $\Lambda(G) = \emptyset \Rightarrow G$ elementary of type (i), $|\Lambda(G)| = 1 \Rightarrow G$ elementary of type (ii), and $|\Lambda(G)| = 2 \Rightarrow G$ elementary of type (iii), so elementary groups are characterised by the size of their limit sets. Indeed

Proposition 3.9 *A Kleinian group G is elementary if and only $|\Lambda(G)| \leq 2$, and non-elementary if and only if $\Lambda(G)$ is infinite.*

Proof. If $\Lambda(G)$ is finite and non-empty then any G orbit in $\Lambda(G)$ is a finite G orbit on $\hat{\mathbf{C}}$ so G is elementary by definition and has $|\Lambda(G)| = 1$ or 2 by the above classification. QED

We state, without proof, the following properties of ordinary and limit sets of Kleinian groups:

Theorem 3.10 *Any Kleinian group G acts discontinuously on $\hat{\mathbf{C}} - \Lambda(G)$. Hence $\hat{\mathbf{C}}$ is the disjoint union of $\Omega(G)$ and $\Lambda(G)$.*

Proposition 3.11 *Let G be a non-elementary Kleinian group. Then any non-empty closed G-invariant subset S of $\hat{\mathbf{C}}$ contains $\Lambda(G)$*

Corollary 3.12 *Let G be a Kleinian group. Then either $\Lambda(G) = \hat{\mathbf{C}}$ or $\Lambda(G)$ has empty interior.*

Corollary 3.13 *Let G be a non-elementary Kleinian group. Then $\Lambda(G)$ is the closure of the set of all fixed points of loxodromic and hyperbolic elements of G.*

Comment. If G has any parabolic elements their fixed points must lie in $\Lambda(G)$, but elliptic elements may have fixed points in either $\Omega(G)$ or $\Lambda(G)$.

Corollary 3.14 *Let G be a non-elementary Kleinian group. Then $\Lambda(G)$ is perfect (and hence, in particular, uncountable).*

Corollary 3.15 *Definitions 1 and 2 for the limit set $\Lambda(G)$ of a non-elementary Kleinian group G are equivalent.*

Proof. We show that the limit set as defined by definition 1 has exactly the same characterising property as that specified by Proposition 3.11 for $\Lambda(G)$ (where we used definition 2). Let S be any closed G-invariant subset of $\hat{\mathbf{C}}$ (note that S must be infinite, since G is non-elementary). Then $C(S)$, the convex hull of S in $\mathcal{H}^3 \cup \hat{\mathbf{C}}$, is also closed and G-invariant. Take any $x \in C(S) \cap \mathcal{H}^3$. Its orbit Gx is contained in $C(S)$ and the accumulation set of this orbit is contained in $C(S) \cap \hat{\mathbf{C}} = S$. Hence S contains the definition 1 limit set of G. QED

The results stated above for ordinary and limit sets of Kleinian groups exhibit a very close analogy with our earlier results on Fatou and Julia sets for rational maps. This raises the question as to whether we can make the *definitions* analogous too. The answer is yes.

Proposition 3.16 *Let G be a Kleinian group. Then $\Omega(G)$ is the largest open subset of $\hat{\mathbf{C}}$ on which the elements of G form an equicontinuous family.*

Proof. Assume G non-elementary (as usual elementary groups can be dealt with on a case by case basis). Then $\Lambda(G)$ contains at least three points (in fact infinitely many) so $\Omega(G)$ is contained in the equicontinuity set by Montel's Theorem. But given any $z \in \Lambda(G)$, by Corollary 3.13 there must be a repelling fixed point of some $g \in G$ arbitrarily close to z, so the family of maps G cannot be equicontinuous at z. QED

We deduce the following two consequences (useful for plotting $\Lambda(G)$).

Theorem 3.17 *Let G be a non-elementary Kleinian group, and U be any open subset of $\hat{\mathbf{C}}$ meeting $\Lambda(G)$. Then*

$$\bigcup_{g \in G} gU = \hat{\mathbf{C}}$$

Proof. The union $\bigcup_{g \in G} gU$ covers all of $\hat{\mathbf{C}}$ except at most two points (else the family G would be equicontinuous on U by Montel's Theorem). But the complement of this union is a finite G-invariant set and therefore empty (since G is non-elementary). QED

The following corollary is immediate.

Corollary 3.18 *Let G be a non-elementary Kleinian group, and U be any open subset of $\hat{\mathbf{C}}$ meeting $\Lambda(G)$. Then*

$$\bigcup_{g \in G} g(U \cap \Lambda(G)) = \Lambda(G)$$

Comments

1. A discrete subgroup of $PSL(2, \mathbf{R})$ is called *Fuchsian.* All our results for Kleinian groups in this chapter have obvious specialisations to the Fuchsian case, with $\mathcal{H}^3$ replaced by $\mathcal{H}^2$, and $\hat{\mathbf{C}}$ replaced by $\hat{\mathbf{R}}$.

2. There are many analogies between rational maps and Kleinian groups; methods and results in each area suggest anoalogous techniques and conjectures in the other, but by no means everything that one might expect to be true or provable turns out to be so. Together the analogies make up the "Sullivan Dictionary" between the two subjects. For the state of this dictionary around five years ago see the book *Holomorphic Dynamics* by Morosawa, Nishimura, Taniguchi and Ueda (CUP 2000), but more entries have been resolved since then: for example the Ahlfors $0-1$ Conjecture (that the limit set of a finitely generated Kleinian group has Lebesgue measure zero if it is not the entire sphere) has recently been proved, and even more recently it has been shown that there exists a polynomial map with Julia set of positive Lebesgue measure.

3.4 Fundamental domains for Kleinian groups, Poincaré's polyhedron theorem

Let G be a Kleinian group, acting on $\mathcal{H}^3$, on $\hat{\mathbf{C}}$, or on $\mathcal{H}^3 \cup \hat{\mathbf{C}}$, and let $\Omega(G)$ be the regular set for the action.

Definition A *fundamental domain* for the action of G on $\Omega(G)$ is a subset F of $\Omega(G)$ such that

$$(i) \qquad \bigcup_{g\in G} g(\bar{F}) = \Omega(G) \quad \text{and}$$

$$(ii) \qquad g(F) \cap h(F) = \emptyset \quad \text{when} \quad g \neq h \quad (g, h \in G)$$

(where in (i), $\bar{F}$ denotes the closure of F).

Thus the images of F *tessellate* $\Omega(G)$ (they cover it without overlapping).

Example The set $\{x+iy : 0 < x < 1\}$ is a fundamental domain for the action of $z \to z+1$ on the complex plane $\mathbf{C}$ (as indeed is the set $\{x+iy : 0 \leq x < 1\}$).

Note The precise definition of the term 'fundamental domain' varies from author to author: some require F to be closed - in which case of course one must modify condition (ii) above to require only that $g(F) \cap h(F)$ be contained in the *boundary* of both $g(F)$ and $h(F)$, rather than it be empty.

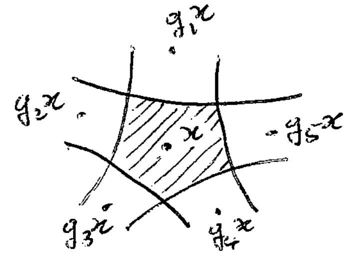
Dirichlet domains

The simplest construction of fundamental domains makes use of a metric. So for the time being we consider an action of G on $\mathcal{H}^3$ (or, if G is Fuchsian, on $\mathcal{H}^2$).

Choose $x \in \mathcal{H}^3$ such that for all $g \in G$ except the identity, $gx \neq x$. (Exercise: show that there are at most a discrete set of points $x \in \mathcal{H}^3$ which do not have this property.) Now for each $g \in G$ define the *half-space*

$$H_g = \{y \in \mathcal{H}^3 : d(y, x) < d(y, gx)\}$$

where $d(y, x)$ denotes the hyperbolic distance from y to x.



Definition The *Dirichlet domain centred at* x is the set

$$D_x = \bigcap_{g \in G - \{I\}} H_g$$

Thus D_x consists of those points of $\mathcal{H}^3$ which are nearer to x than they are to any gx $(g \in G - \{I\})$.

This construction was introduced by Dirichlet in the 1850's for the study of Euclidean groups, and later adapted by Poincaré for the hyperbolic case.

Proposition 3.19 *For any Kleinian group G, a Dirichlet domain D_x is a fundamental domain for the action of G on $\mathcal{H}^3$.*

Proof. We must prove that D_x satisfies conditions (i) and (ii) of the definition of a fundamental domain. We first observe that

$$g(D_x) = \{y : d(y, gx) < d(y, hx) \quad \forall h \in G - \{g\}\}$$

since

$$y \in g(D_x) \Leftrightarrow g^{-1}y \in D_x \Leftrightarrow d(g^{-1}y, x) < d(g^{-1}y, kx) \Leftrightarrow d(y, gx) < d(y, gkx) \; \forall k \in G - \{I\}$$

Now take any $y \in \mathcal{H}^3$. Take $g \in G$ (not necessarily unique) such that $d(y, gx)$ is minimal. Then $y \in g(\bar{D}_x)$ so property (i) holds. Moreover it is clear that $g(D_x) \cap h(D_x) = \emptyset$ if $g \neq h$ so property (ii) holds too. QED

Recall that a subset $X \subset \mathcal{H}^3$ is said to be *convex* if given any $x, y \in X$ the segment of geodesic joining x to y is entirely contained in X.

Proposition 3.20 *A Dirichlet domain D_x for a Kleinian group G is convex and locally finite (i.e. each compact subset K of $\mathcal{H}^3$ meets only finitely many $g(D_x)$).*

Proof. Convexity is obvious since D_x is defined to be an intersection of half-spaces, each of which is convex.

For local finiteness, take the Poincaré disc model of $\mathcal{H}^3$ and without loss of generality take x to be the origin and K to be the closed ball with centre the origin and (hyperbolic) radius ρ. We claim that if g is any element of G such that $gD_0 \cap K$ is non-empty then $d(0, g0) \leq 2\rho$, which will prove local finiteness since G, being discrete, contains only finitely many elements with $d(0, g0) \leq 2\rho$ (else the orbit of 0 would have an accumulation point in $\mathcal{H}^3$, contradicting discontinuity of the action of G there). To prove the claim, take any $y \in gD_0 \cap K$; then $d(0, y) \leq \rho$ (since $y \in K$) and $d(g0, y) \leq d(0, y)$ (since $y \in gD_0$) so $d(0, g0) \leq \rho + \rho = 2\rho$. QED

Definition A convex region P obtained as the intersection of countably many half spaces H_j in $\mathcal{H}^3$, with the property that any compact subset of P meets only finitely many of the hyperplanes ∂H_j is called a *polyhedron* (and a subset of $\mathcal{H}^2$ with the analogous property is called a *polygon*).

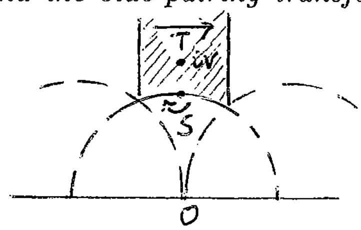
Thus Proposition 3.20 says that a Dirichlet domain is a polyhedron. Note that the proposition does not say that D_x has only *finitely many faces*, at least it only says this when D_x is *compact*. When D_x has finitely many faces (for some x) we say that G is *geometrically finite*.

Now consider any point y on the boundary of D_x, so y is on the boundary of H_g for one of the half-spaces defining D_x, in other words $d(y,x) = d(y,gx)$ for some $g \in G$. Then

$$d(g^{-1}y, g^{-1}x) = d(y,x) = d(y,gx) = d(g^{-1}y, x)$$

so $g^{-1}y$ also lies in the boundary of D_x. Thus each face of D_x is carried to another face of D_x by an appropriate element of G. We call these elements *side-pairing transformations.* For an example consider the action of $PSL(2,\mathbf{Z})$ on $\mathcal{H}^2$.

Proposition 3.21 *Let $G = PSL(2,\mathbf{Z})$ act on the complex upper half-plane in the usual way. Then for any point iv on the imaginary axis, with $v > 1$, the Dirichlet domain is the region illustrated below, and the side-pairing transformations are $T : z \to z+1$, $S : z \to -1/z$.*



Proof. Let $P = \{z : |Re(z)| < 1/2, |z| > 1\}$ (the region illustrated) and let D_{iv} denote the Dirichlet region, centred on iv, for G. We first observe that $D_{iv} \subset P$ since P is the set of points nearer to iv than to $iv-1, iv+1$ and i/v (in the hyperbolic metric).

It remains to show that there no are points z' in P which are nearer to $g(iv)$ than to iv for some other $g \in G$. Suppose there is such a z'. Then $z' \in h(D_{iv})$ for some $h \in G$ (since the translates of D_{iv} cover the upper half-plane, by Proposition 5.1). Let $z = h^{-1}z' \in D_{iv}$. Now both z and hz lie in P, and we obtain a contradiction as follows. Suppose

$$h(z) = \frac{az+b}{cz+d}$$

(with the matrix normalised to have determinant one). Then by an easy exercise

$$Im(h(z)) = \frac{Im(z)}{|cz+d|^2}$$

But

$$|cz+d|^2 = c^2|z|^2 + 2Re(z)cd + d^2 > c^2 + d^2 - |cd| = (|c| - |d|)^2 + |cd|$$

(since $|z| > 1$ and $|Re(z)| < 1/2$). However $(|c|-|d|)^2 + |cd|$ is a positive integer (since $c = d = 0$ is ruled out by $ad - bc = 1$). So $|cz+d| > 1$, and hence $Im(h(z)) > Im(z)$. But the same reasoning applied to $h(z)$ and h^{-1} in place of z and h yields $Im(z) < Im(h(z))$ and hence a contradiction. QED

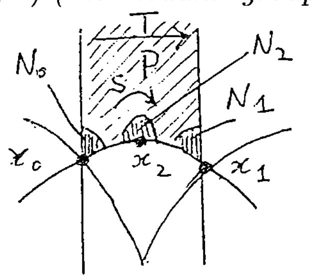
Poincaré's Polyhedron Theorem

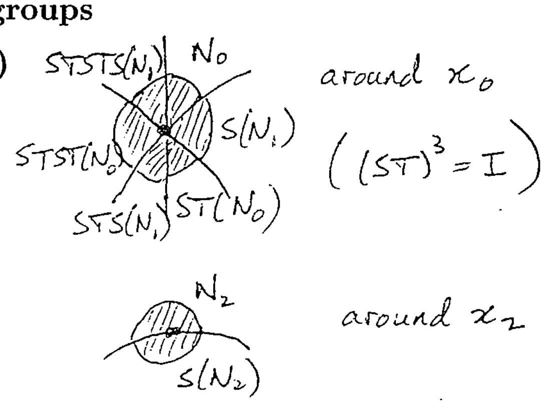
We have seen that given a Kleinian group G, Dirichlet's construction allows us to find a fundamental domain on which G acts by side-pairing transformations. Poincaré's Polyhedron Theorem takes us in the opposite direction: given a convex polyhedron in $\mathcal{H}^3$ (or polygon in $\mathcal{H}^2$) and a set of side-pairing transformations for that polyhedron, it gives us necessary and sufficient conditions for the group generated by those transformations to be discrete (i.e. Kleinian) and for the given polyhedron to be a fundamental domain for the group action. (Note that for a polyhedron, the 'side-pairings' identify *faces* of the polyhedron.) The key condition is that the translates of the polyhedron under the group generated by the side pairings should fit together "without overlaps" along each edge and around each vertex. Poincaré's theorem then yields a presentation of the group, with the side-pairing transformations as generators, and a relation for each *edge*, in the 3-dimensional polyhedron case. In the 2-dimensional polygon case the side-pairings identify *edges* and we have a relation for each *vertex*. We refer the reader to Beardon's book on discrete groups, or to Ratcliffe or Maskit, as the precise conditions, though conceptually straightforward, are a little cumbersome to state. Our main concern here will be to understand examples.

3.5 Examples of Fuchsian and Kleinian groups

Examples in $PSL(2,\mathbf{R})$ (Fuchsian groups)

1. $PSL(2,\mathbf{Z})$ *(the modular group)*

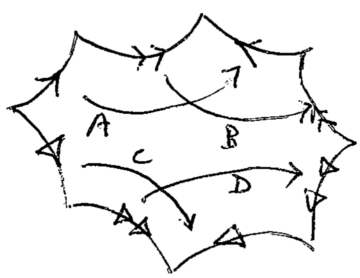




Around x_1 the picture is just that around x_0, conjugated by T. The vertex $y_0 = \infty$ is ideal, and T is parabolic ($z \to z+1$). Hence P is a fundamental domain for $PSL(2,\mathbf{Z})$, as we have already proved earlier. Poincaré's Polygon Theorem tells us that

$$PSL(2,\mathbf{Z}) = < S, T : S^2 = I, (ST)^3 = I >$$

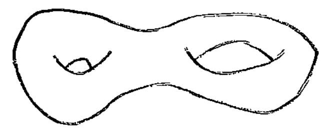
2. *Surface groups*



In the picture P is a regular octagon with vertex angles all $\pi/4$. (To find such an octagon in the Poincaré disc model, just take a small regular octagon centred at the origin and blow it up steadily in size unit the angles are $\pi/4$: this case must occur, by continuity, since in the limiting case when all vertices are ideal the angles are 0). Let A, B, C, D be the side pairings shown. Then P is a fundamental domain for the group

$$G =< A, B, C, D : [A, B][C, D] = I >$$

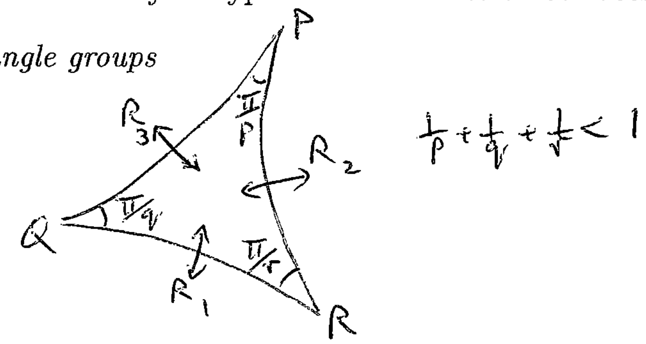
(where $[A, B][C, D] = ABA^{-1}B^{-1}CDC^{-1}D^{-1}$). Note that $\mathcal{H}^2/G$ is a a surface of genus two:



Higher genus surfaces may be obtained similarly.

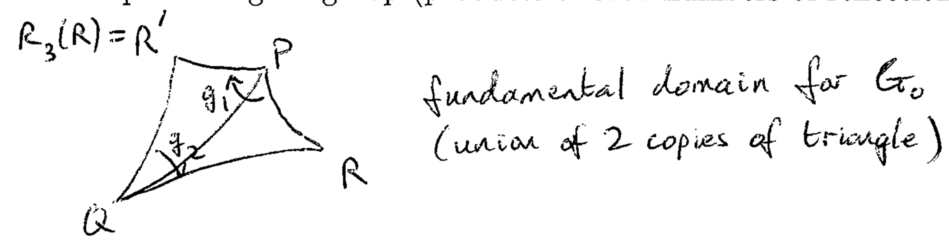
Comment The octagon need not be regular: all that is really needed is that the angles add up to 2π and that the sides paired be of the same length. This is the beginning of the Teichmüller theory of hyperbolic stuctures on surfaces.

3. *Triangle groups*



We can always draw such a triangle in $\mathcal{H}^2$ by taking a small Euclidean triangle at the origin in the Poincaré disc model and gradually enlarging it until the angles are those desired. The (hyperbolic) area of such a triangle is π minus the angle sum.

Now let G be the group generated by reflections in the sides of the triangles, and let G_0 be its orientation-preserving subgroup (products of even numbers of reflections).

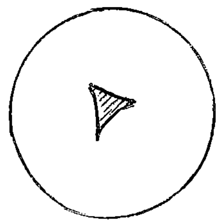


G_0 has generators $g_1 = R_2R_3$ and $g_2 = R_3R_1$. By Poincaré's Theorem G_0 is discrete, the quadrilateral shown is a fundamental for G_0, and a presentation for G_0 is

$$G_0 =< g_1, g_2 : g_1^p = g_2^q = (g_1g_2)^r = I >$$

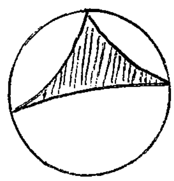
(Note that if $1/p + 1/q + 1/r > 1$ we can construct a *spherical triangle* and the group G_0 is then *finite*.)

4. *Limit sets of triangle and truncated triangle groups*

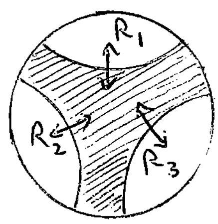


When the fundamental polygon for G is compact, the limit set of G is the entire boundary circle S^1 of the Poincaré disc (the translates of P get smaller and smaller in the Euclidean metric as we move towards the boundary circle, so the orbit of of any point inside the disc accumulates everywhere on S^1).

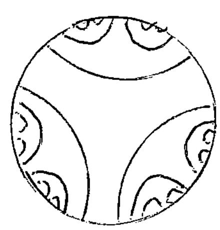
When the fundamental domain has ideal vertices (that is to say vertices on the *boundary* of hyperbolic space) the limit set remains the entire circle:



However we can go further and take for example a 'truncated triangle' for our polygon P:



As before let G be the group generated by reflections R_1, R_2, R_3, and G_0 be the orientation-preserving subgroup (generated by R_2R_3, R_3R_1). Now R_2R_3 is hyperbolic and the 'gap' between its fixed points is in $\Omega(G_0) \subset S^1$. hence $\Lambda(G_0) \neq S^1$, so $\Lambda(G)$ has empty interior in S^1. hence $\Lambda(G)$ is totally disconnected, But $\Lambda(G)$ is infinite, perfect, closed and bounded. Hence $\Lambda(G)$ is a Cantor set.

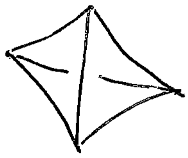


Note that G_0 is freely generated by R_2R_3 and R_3R_1: there are no vertices so no relations.

Examples in $PSL(2, \mathbf{C})$ (Kleinian groups)

1. *Tetrahedron groups*

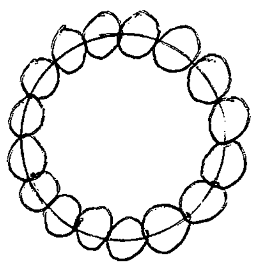
Our 'polygon' now becomes a tetrahedron in $\mathcal{H}^3$ rather than a triangle in $\mathcal{H}^2$, and we consider the group G generated by reflections in its faces, and the orientation preserving subgroup G_0.



A tetrahedron in $\mathcal{H}^3$ is determined by its six *dihedral angles* (the angles between adjacent faces). To satisfy the conditions of Poincaré's Theorem we require them all to be of the form π/n with n integer. A vertex inside $\mathcal{H}^3$ must have $1/p_1 + 1/p_2 + 1/p_3 > 1$, an ideal vertex must have $1/p_1+1/p_2+1/p_3 = 1$, and a truncated vertex must have $1/p_1+1/p_2+1/p_3 < 1$. Where there is a truncated vertex the tetrahedron must meet the boundary of $\mathcal{H}^3$ in a $\pi/p_1, \pi/p_2, \pi/p_3$ triangle.

One can show that all combinations of dihedral angles are actually realised by tetrahedra or truncated tetrahedra. If all the vertices are internal or ideal then $\Lambda(G) = \hat{\mathbf{C}}$. If one or more vertices is truncated then $\Lambda(G)$ is a circle-packing (we get a circle as limit set for the triangle group around the truncated vertex, and then other elements of G move this circle around).

2. *'Strings of beads'*



Here $C_1, ..., C_n$ are circles in $\bar{C}$, each of the same size, touching the circle on each side and orthogonal to the circle S_1. Let R_m denote inversion in C_m, and extend R_m to a reflection in the hemisphere H_m spanning C_m in $\mathcal{H}^3$.

Now, by Poincaré's Theorem, the part of $\mathcal{H}^3$ remaining after 'scooping out' all the hemispheres is a fundamental domain for the action of $G =< R_1, ..., R_n >$ and the only relations are $R_m^2 = I >$.

Note that the limit set here is S^1, but that if we pull the circles C_m apart the limit set becomes a Cantor set, and that if we perturb the sizes and positions of the circles C_m, but keeping them touching adjacent circles, the limit set becomes a *quasicircle* (a fractal homeomorphic to a circle). Going up in dimension an analogous construction can be used to obtain a group having limit set a wildly embedded circle in S^3.

LECTURE 4. QUADRATIC MAPS AND THE MANDELBROT SET

4.1 The Mandelbrot set and its connectivity

Proposition 4.1 Every quadratic map $f(z) = \alpha z^2 + \beta z + \gamma$ with $\alpha \neq 0$ is conjugate to $q_c(z) = z^2 + c$ for a unique c.

Proof The conjugacy h must send ∞ to itself, and hence have the form $h(z) = kz + l$.

$$hf(z) = k(\alpha z^2 + \beta z + \gamma) + l \qquad q_c h(z) = (kz + l)^2 + c$$

These are equal (for all z) if and only if $k\alpha = k^2, k\beta = 2kl$ and $k\gamma + l = l^2 + c$. Thus we must have $k = \alpha, l = \beta/2$ and $c = \alpha\gamma + \beta/2 - \beta^2/4$. QED

Another useful parametrisation of the quadratic maps is given by the *logistic family*

$$p_\lambda(z) = \lambda z(1 - z)$$

Clearly p_λ is conjugate to q_c if and only if $c = \lambda/2 - \lambda^2/4$ (by Proposition 4.1).

The q_c parametrisation is more convenient when we are dealing with critical points, and the p_λ parametrisation is more convenient when we are dealing with fixed points and their multipliers. Note that q_c has critical points $0, \infty$, the latter a superattracting fised point, and p_λ has fixed points 0 and $1 - 1/\lambda$, with multipliers λ and $2 - \lambda$ respectively.

Definition

The *Mandelbrot set* is the subset of parameter space defined by

$$M = \{c : J(q_c) \text{ connected}\} \subset \mathbf{C}$$

Theorem 4.2 *M is the set of values of the parameter c such that the orbit $q_c^n(0)$ of the critical point 0 does not tend to the point ∞*

Proof If the orbit of 0 does not tend to ∞ then there is no obstruction to extending the Böttcher coordinate to the whole of the basin B_∞ of attraction of ∞. Hence B_∞ is homeomorphic to the open unit disc and its complement $\hat{\mathbf{C}} - B_\infty$ is connected as is their common boundary ∂B_∞. But ∂B_∞ is closed and completely invariant, and cannot contain any points of the Fatou set (since any point in ∂B_∞ has bounded orbits, yet arbitrarily close to it are points with orbits going to ∞): hence ∂B_∞ is the Julia set $J(q_c)$.

Conversely, if the orbit of 0 does go to ∞ then $J(q_c)$ is totally disconnected (a Cantor set) by the argument sketched earlier for the example $|c|$ large. QED

Definition The *filled Julia set* of q_c is $K(q_c) = \{z : q_c^n(0) \not\to \infty\}$

(Note that if $c \notin M, K(q_c) = J(q_c)$ =Cantor set.)

Theorem 4.3 (Douady and Hubbard 1982) *The Mandelbrot set M is connected*

Proof In fact they proved a much stronger result, that there is a conformal bijection between the complement $\hat{\mathbf{C}} - M$ of the Mandelbrot set and the complement $\hat{\mathbf{C}} - D$ of the open unit disc. It is an immediate consequence of this that M is connected.

When $c \in M$, the Böttcher coordinate defines a conformal bijection

$$\phi_c : \hat{\mathbf{C}} - K(q_c) \to \hat{\mathbf{C}} - D$$

$$\phi_c(z_0) = z_0(1 + \frac{c}{z_0^2})^{1/2}(1 + \frac{c}{z_1^2})^{1/4}(1 + \frac{c}{z_2^2})^{1/8}...$$

(conjugating q_c to $z \to z^2$). When $c \notin M$ the map ϕ_c though not defined on the whole of the complement of K_c is nevertheless defined on a neighbourhood of ∞ and as far as the critical value c of q_c. Define

$$\Psi : \hat{\mathbf{C}} - M \to \hat{\mathbf{C}} - D$$

$$\Psi(c) = \phi_c(c)$$

This is a conformal bijection (see Douady and Hubbard, Comptes Rendues 1982, for more details). QED

Conjecture ('MLC') *M is locally connected*

If M is locally connected then by a theorem of Carathéodory the map Ψ^{-1} extends to a continuous map from the boundary of $\hat{\mathbf{C}} - D$ (a circle) onto the boundary ∂M of the Mandelbrot set. This would give us a purely combinatorial description of ∂M and many open questions concerning M would be resolved.

Definition A component of the interior of M is said to be *hyperbolic* if for every c in the component q_c has an attracting or superattracting periodic orbit.

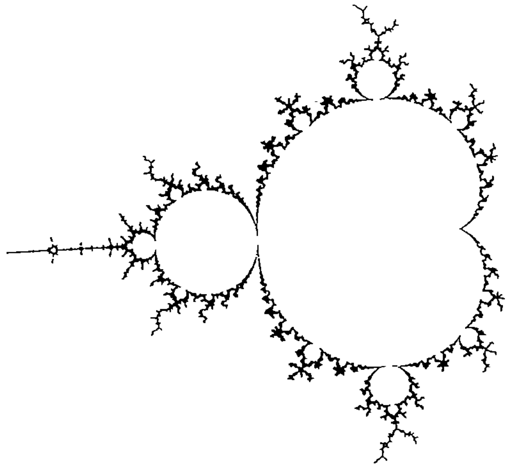
Conjecture ('Hyperbolicity is dense') *Every component of the interior of M is hyperbolic*

Douady and Hubbard showed in their 1985 Orsay lecture notes that 'MLC' implies 'Hyperbolicity is dense'.

Both conjectures seem to be very difficult to resolve. Over the last 15 years there has been a great deal of work on them. The set of points of ∂M at which local connectivity is known to hold has been steadily increased: Yoccoz has proved it for 'all but infinitely renormalizable points' and Lyubich has extended this result to certain of the remaining points. Most experts seem to believe that MLC should be true, but it is known that the analogous set for cubics in place of quadratics is *not* locally connected (Lavaurs, Milnor), and it is known that there exist quadratic maps q_c having non-locally-connected Julia sets. As far as 'Hyperbolicity is dense' is concerned, this has been proved for components of M meeting the real axis (Lyubich, McMullen, Swiatek: see McMullen's 1994 book 'Complex Dynamics and Renormalization') but the general question is still unresolved. Another recent development is Shishikura's proof (1994) that the boundary ∂M of the Mandelbrot set has Hausdorff dimension 2.

4.2 The geography of the Mandelbrot set

We examine some of the more prominent features of M.



Let $M_0 = \{c : q_c \textit{ has an attracting (or superattracting) fixed point } \}$
$= \{c : J(q_c) \textit{ is a (topological) circle}\}$

Lemma 4.4 $M_0 = \{c : c = \lambda/2 - \lambda^2/4 \textit{ for some } \lambda \textit{ with } |\lambda| < 1\}$

Proof Consider the logistic map p_λ. The multipliers of its fixed points are $\lambda, 2-\lambda$. Hence
$M_0 = \{c : c = \lambda/2 - \lambda^2/4 \textit{ for some } \lambda \textit{ with } |\lambda| < 1 \textit{ or } |2-\lambda| < 1\}$
But $\lambda/2 - \lambda^2/4 = (2-\lambda)/2 - (2-\lambda)^2/4$. QED

Thus M_0 is a *cardioid* (with a boundary that is smooth except at the cusp $c = 1/4$). Note that there is a bijection between points of M_0 and values of λ such that $|\lambda| < 1$. Thus M_0 is parametrised by the multiplier of the fixed point of q_c.

The intersection of M with the real axis

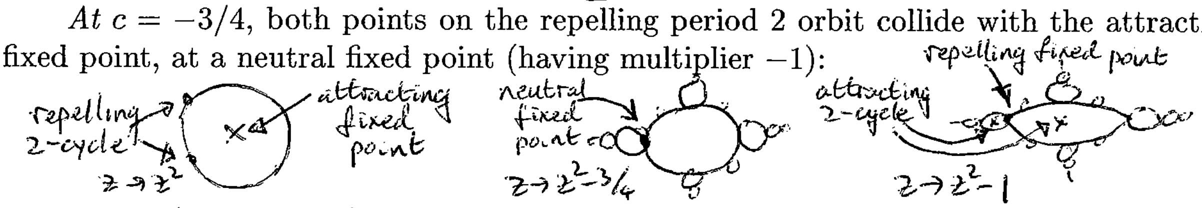
We consider how the behaviour of q_c varies as we vary the parameter c along the real axis.

For $c > 1/4$, $J(q_c)$ is a Cantor set (it is an easy exercise to show that the orbit of 0 under q_c tends to ∞).

At $c = 1/4$, there is a neutral fixed point $z = 1/2$, with multiplier 1.

For $-3/4 < c < 1/4$ q_c has an attractive fixed point and $J(q_c)$ is a (topological, indeed *quasi-conformal*) circle, with dynamics conjugate to that of the shift. In particular $J(q_c)$ contains a dense set of repelling periodic orbits.

At $c = -3/4$, both points on the repelling period 2 orbit collide with the attracting fixed point, at a neutral fixed point (having multiplier -1):



For $-5/4 < c < -3/4$, q_c has an attractive period 2 orbit, and the topology of $J(q_c)$ is the same as that (plotted earlier) for the special (superattractive) case $c = -1$.

We digress briefly to justify the bounds $-5/4 < c < -3/4$:

Lemma 4.5 *q_c has an attracting period 2 orbit if and only if* $|1+c| < 1/4$

Proof The points of period 1 or 2 are the solutions of $q_c^2(z) = z$. Expanding $q_c^2(z) - z$ we have

$$q_c(q_c(z)) - z = ((z^2+c)^2 + c - z = (z^2 - z + c)(z^2 + z + 1 + c) = (z-\alpha)(z-\beta)(z-u)(z-v)$$

where α, β are the fixed points and u, v is the period 2 cycle. The multiplier of the period 2 cycle is $q_c'(u)q_c'(v) = 4uv = 4(1+c)$. The period 2 cycle is attracting if and only if this has modulus less than 1. QED.

Returning to our journey in parameter space along the real axis, for $-2 < c < -5/4$, as c decreases through this range, we have a sequence of period doublings until we reach the Feigenbaum point. This is followed by the whole Milnor/Thurston sequence of periods, in particular containing the Sarkovskii sequence, the most prominent component of $int(M)$ along the axis being that corresponding to a period three attracting orbit, and we finish at $c = -2$ where the Julia set is the real interval $[-2, +2]$ (and q_c is semi-conjugate to $z \to z^2$: see the exercise early on in these notes). For $c < -2$, it is easily proved that the orbit of the critical point 0 tends to ∞ and hence the Julia set is again a Cantor set.

The behaviour for c at different points along the real axis is is no surprise since the quadratic family is conjugate to the logistic family. However with c complex we can now leave the main cardioid M_0 at other points than just $c = -3/4$. When c is on the boundary of M_0 at the point where $\lambda = e^{2\pi i p/q}$, q_c has a neutral periodic point with this as multiplier, and when c passes into the adjoining component q_c has an attracting period q orbit. There are then further bifurcations as we pass along a path through different components of M_0. We shall have more to say about this in the next section.

Exercise Compute the values of c where q_c has a superattractive period three orbit (that is, where the point 0 has period three).

4.3 Internal and external rays: the 'devil's staircase'

When $c \in M$, for any $\theta \in [0,1)$, the radial line $arg(z) = 2\pi\theta$ on $\hat{\mathbf{C}} - D$ (where D is the unit disc) maps under the inverse ϕ_c^{-1} of the Böttcher map to the *external ray* $\mathcal{R}_\theta$ of argument $2\pi\theta$ on $\hat{\mathbf{C}} - K(q_c)$.

Similarly, in the parameter plane, the radial line $arg(z) = 2\pi\theta$ on $\hat{\mathbf{C}} - D$ maps under the inverse Ψ^{-1} of the Douady-Hubbard map to the *external ray* $\mathcal{R}_\theta$ of argument $2\pi\theta$ on $\hat{\mathbf{C}} - M$.

The combinatorics of these rays can tell us a great deal about the structure of M.

A ray is said to *land*, if it accumulates at a unique point of $J(q_c)$ (in the dynamical case) or ∂M (in the parameter case). If $J(q_c)$ (or ∂M respectively) is locally connected then all external rays land (by Carathéodory's criterion). Unfortunately there are examples where $J(q_c)$ is known not to be locally connected, and where certain external rays do not

But supposing we have found this orbit $A_{p/q}$, how are we to know which of its points are the special points $\theta_\pm(p/q)$? This turns out to be very straightforward.

Lemma 4.7 *Any ordered orbit of $t \to 2t$ on the circle $\mathbf{R}/\mathbf{Z}$ is contained in a semicircle*

Proof Since $t \to 2t$ doubles distance, any three points on the circle have images in the same order around the circle if and only the three original points lie in a common semi-circle. QED

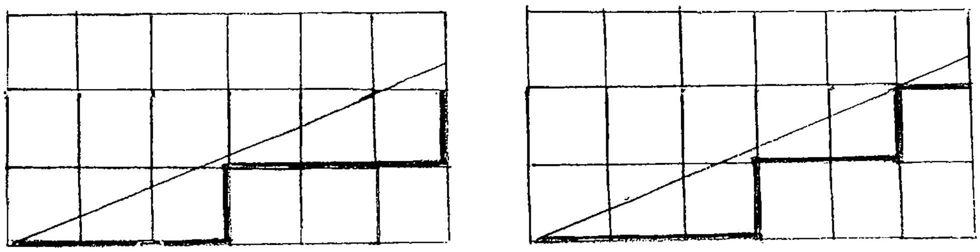
As a consequence it makes sense to refer to the *least* and *greatest* points of the orbit $A_{p/q}$. We identify the points $\theta_\pm(p/q)$ by observing that the dynamical picture requires that the least point of $A_{p/q}$ be $(\theta_+(p/q))/2$ and the greatest be $(\theta_-(p/q))/2 + 1/2$ (see the picture above for the case $p/q = 1/3$: the inverse image of the component of $int(K(q_c))$ containing the critical value c is that containing the critical point 0).

Algorithm for $\theta_\pm(p/q)$

There is a simple algorithm constructing the binary sequence of each of $\theta_+(p/q)$ and $\theta_-(p/q)$:

Draw a line of slope p/q, through the origin in $\mathbf{R}^2$. To construct $\theta_-(p/q)$, take the integer 'staircase' lying just below this line, but not touching it, and starting at the point $(1,0)$ write 1 for each horizontal step which is followed by a vertical step, 0 for a horizontal step followed by another horizontal one. To construct $\theta_+(p/q)$ do the same with the staircase touching the line.

Example $p/q = 2/5$



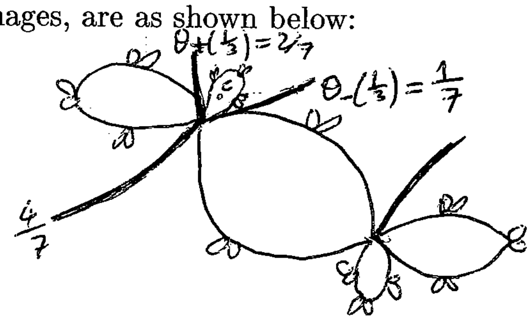
$\theta_-(2/5) = .\overline{01001} = 9/31$ $\quad\quad$ $\theta_+(2/5) = .\overline{01010} = 10/31$

(For a justification of this algorithm see Bullett and Sentenac.)

Every point on ∂M_0 at the end of internal rays of *irrational* argument ν is the landing point of a single external ray, of argument $\theta_\nu = lim_{p/q \to \nu}\theta_\pm(p/q)$. The assigment of internal angles to external angles as we make a circuit of the boundary of the cardioid M_0 has graph a 'devil's staircase' (see the graph on the next page).

land; moreover the conjecture 'MLC' is still unproved so we cannot be sure that all external rays in the parameter space land. However it has been proved by Douady and Hubbard that for all $c \in M - int(M_0)$ (where M_0 is the cardioid) the fixed point α of q_c (the fixed point in the dynamical plane which does not correspond to the external ray of argument zero) is the landing point of a finite set of external rays, and that in the parameter plane all points on the boundaries of hyperbolic components of $int(M)$ are landing points of external rays: we shall restrict ourselves to these rays in the discussion below.

We start by considering the external rays which land on the boundary of the main cardioid, M_0. Recall that M_0 is itself parametrised by the unit disc and we can therfore define *internal rays* inside M_0. The *internal ray* of argument ν is the set of values of $c \in M_0$ for which the multiplier of the fixed point of q_c has argument $2\pi\nu$. Consider the end point on ∂M_0 of the internal ray of argument $\nu = 1/3$. This is the value of c for which the fixed point α of q_c has multiplier $e^{2\pi i/3}$ (this c lies at the top of the cardioid: it is where the first period-tripling occurs). The external rays in the dynamical plane landing at α, and one of their inverse images, are as shown below:

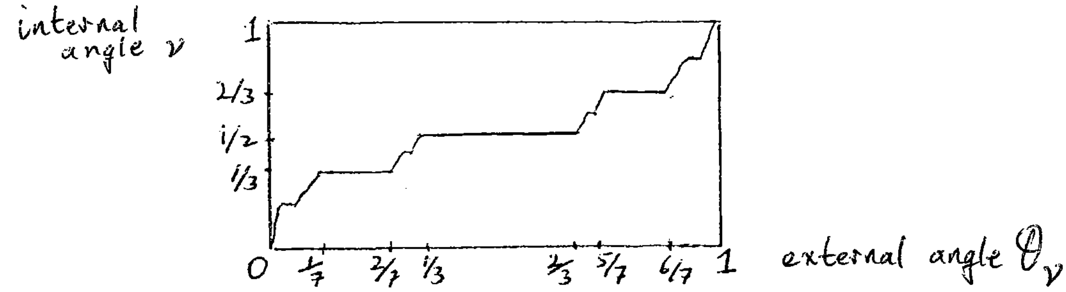


Note that we can pick out two particular rays, which together enclose the component of $int(K(q_c))$ containing the critical value. These we have labelled $\theta_-(1/3)$ and $\theta_+(1/3)$. It can be shown that in the parameter space the corresponding external rays with the same arguments, $\theta_-(1/3)$ and $\theta_+(1/3)$, land at c (an example what Douady calls 'ploughing in the dynamical plane but harvesting in the parameter plane').

More generally, for c at the end of each internal ray in M_0 of rational argument p/q, the map q_c has a (neutral) fixed point α of rotation number p/q and we can pick out the pair of external rays enclosing the component of $int(K(q_c))$ containing the critical value c. How do we compute the values of $\theta_-(p/q)$ and $\theta_+(p/q)$? Since α is a fixed point of rotation number p/q there are necessarily q external rays landing at α and the effect of q_c on these rays is to permute them in cyclic order. But the action of q_c on arguments of rays is simply that of $t \to 2t$ (mod $\mathbf{Z}$), so our search for candidates for $\theta_\pm(p/q)$ is reduced to a search for finite orbits of $t \to 2t$ on the unit circle $\mathbf{R}/\mathbf{Z}$, arranged in the same order around the circle as an orbit of a rigid rotation through $2\pi p/q$. This is a purely combinatorial question and was answered (though in a slightly different context) by Morse and Hedlund in their pioneering work on symbolic dynamics in the 1930's:

Theorem 4.6 *For each rational p/q there is a unique finite forward invariant orbit $A_{p/q}$ of $t \to 2t$ of rotation number p/q on the circle $\mathbf{R}/\mathbf{Z}$.*

(For a proof of this and other results concerning order-preserving orbits of the shift, see Bullett and Sentenac, Math. Proc. Cam. Phil. Soc. 1994.)



We draw the graph this way round (rather than that assigning external angles to internal angles) in order to have a *continuous* function. It is not difficult to prove that the horizontal steps in the graph above have total length 1 (see Bullett and Sentenac). A 'devil's staircase' is the graph of a continuous function that is constant on a set of full measure without being globally constant.

Remark For all c in the 'p/q-limb' of the Mandelbrot set (the subset of M attached to the main cardioid M_0 at the point of ∂M_0 of internal angle p/q) the α-fixed point of q_c is the landing point of q external rays permuted cyclically with rotation number p/q by q_c. (We say that the α-fixed point of q_c has *combinatorial rotation number* p/q). The devil's staircase can be thought of as the assignment of this combinatorial rotation number as c circles once around the boundary ∂M of the Mandelbrot set (at least it can be though of this way if ∂M is indeed locally connected).

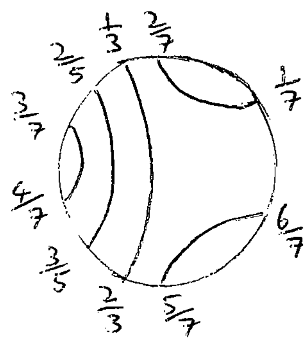
4.4 The combinatorial Mandelbrot set

We sketch an algorithm due to Lavaurs (Comptes Rendues 1986) which, if ∂M is locally connected, gives M as the quotient of the unit disc by an equivalence relation defined via a *lamination*.

Lavaurs' Algorithm

1. Connect 1/3 to 2/3 (on ∂D) by an arc in D.

2. Connect pairs of points p/q in order of increasing q, and increasing p for each q, each time connecting a point to the first subsequent point possible without crossing arcs already constructed.



The (combinatorial) Mandelbrot set is now obtained by shrinking each of the arcs to points. One can also obtain combinatorial models of filled Julia sets by shrinking appropriate laminations of the disc.

4.5 Kneading sequences and internal addresses in the Mandelbrot set

We have dealt with the external rays landing at points on the boundary of the main cardioid M_0. What can we say about the arguments of external rays landing at other points on the boundary of M ? A first step is to consider those landing on the boundary of a component of $int(M)$ immediately adjacent to M_0, say that corresponding to rotation number p/q. This component (which we shall label $M_{p/q}$) has the property that corresponding maps q_c each have an attractive period q orbit. We can parametrise $M_{p/q}$ by the multiplier of this orbit and hence define internal rays inside $M_{p/q}$ in just the same way as we did for M_0. The r/s internal ray in $M_{p/q}$ is the landing point of external rays $\theta_\pm(p/q, r/s)$ obtained from $\theta_\pm(r/s)$ by replacing the digit 0 by the repeating block (of length q) from $\theta_-(p/q)$ and the digit 1 by the repeating block from $\theta_+(p/q)$.

Example

$$\theta_-(1/3, 1/2) = .\overline{001010} \qquad \theta_+(1/3, 1/2) = .\overline{010001}$$

By repeating the same process (which is known as 'tuning') we can compute the arguments of external rays landing on the boundary of any component which is accessible from M_0 by a finite number of boundary crossings. But there are of course components of $int(M)$ which are much further away than this from M_0: for example all components beyond the Feigenbaum point on the real axis are an infinite number of boundary crossings away from M_0. We conclude this section with a brief discussion of the method of Lau and Schleicher assigning 'internal addresses' to all the hyperbolic components of $int(M)$ (SUNY Stony Brook preprint 1994): a related method was developed independently and earlier by Penrose (Warwick thesis 1990).

Internal addresses

Any hyperbolic component $\mathcal{A}$ of $int(M)$ is accessible from M_0 by a path in M passing through a unique (though possibly infinite) sequence of components. We write down the sequence of 'least periods corresponding to components further down the path': this sequence of least periods constitutes the *internal address* of $\mathcal{A}$.

Examples The internal address of the Feigenbaum point is $(1, 2, 4, 8, ...)$. The internal address of the period 3 component on the real axis is $(1, 2, 3)$.

A further refinement is the *angled internal address* where one attaches to each 'least period' the internal angle at which one departs from the corresponding component. Lau and Schleicher showed that angled internal addresses determine components uniquely, and they gave algorithms for converting internal addresses into *kneading sequences* and vice versa.

Kneading sequences

The *kneading sequence* $K(\theta)$ of an external ray of argument θ is the sequence obtained from the orbit of θ (under the map $t \to 2t$ on the circle) by writing '1' when the point lies in the semicircle $(\theta/2, \theta/2 + 1/2)$ and '0' when it lies in the semicircle $(\theta/2 + 1/2, \theta/2)$ (there are conventions for what happens on the boundaries, but these need not concern us here). By the *kneading sequence* of a hyperbolic component $\mathcal{A}$ we mean the limit of $K(\theta)$ as $\theta \to \theta_-(\mathcal{A})$ *downwards*, which can be shown to be equal to the limit of $K(\theta)$ as $\theta \to \theta_+(\mathcal{A})$ *upwards*. For example each of our $M_{p/q}$ attached to M_0 has kneading sequence $\overline{111...10}$ ($(q-1)$ 1's followed by a 0).

Converting internal addresses to kneading sequences and vice versa

Let $\mathcal{A}$ be a hyperbolic component with period n and internal address $(1 = n_1, n_2, ..., n_k = n)$. Then $K(\mathcal{A})$ starts with a 1, and the first n_{i+1} entries of $K(\mathcal{A})$ are obtained by continuing the first n_i entries periodically then changing the n_{i+1}st entry. $K(\mathcal{A})$ is periodic with period n. Conversely, given $K(\mathcal{A})$ recover the internal address as follows. Start with $K = \bar{1}$ and compare it with $K(\mathcal{A})$; the position n of the first difference will be the next entry in the internal address. Repeat the comparison, taking K as the periodic continuation of the first n entries of $K(\mathcal{A})$, and continue until the period of $\mathcal{A}$ is reached.

Examples

1. The period 3 component with internal address $(1, 2, 3)$ has kneading sequence $\overline{100}$.
2. The Feigenbaum point has external angle 0110100110010110... (the first digit is '0', the second digit it the opposite to the first, the third and fourth digits are opposite to the first and second,...etc.). It follows (exercise) that the kneading sequence of the Feigenbaum point is 1011101010111011..., and hence that the internal address of the Feigenbaum point is $(1, 2, 4, 8, ...)$.

4.6 Further topics: polynomial-like maps and their applications (a brief introduction)

One of the central tools in the study of both rational maps and Kleinian groups is *quasiconformal deformation theory*, introduced in the 1960s by Ahlfors and Bers.

A homoeomorphism f between open domains in $\mathbf{C}$ is said to be K-quasiconformal if locally it has distributional derivatives in L^2, and its complex dilatation

$$\mu(z) = \frac{\partial f/\partial \bar{z}}{\partial f/\partial z}$$

satisfies $|\mu| \leq (K-1)/(K+1)$ almost everywhere - in other words small circles are mapped to small ellipses of bounded ellipticity. A key result is:

The "Measurable Riemann Mapping Theorem": *Any measurable ellipse field μ on $\mathbf{C}$ with $||\mu||_\infty$ is realisable by a quasiconformal homeomorphism $\phi : \mathbf{C} \to \mathbf{C}$ which has complex dilatation μ. This ϕ is unique if it fixes* 0 *and* 1..

Sullivan used quasiconformal deformation theory in his proof of the "no wandering domains theorem". Douady and Hubbard used in in their theory of "polynomial-like mappings", which explained the appearance of *self-similarity* in Julia sets and in the Mandelbrot set. Their theory played has played a central role in the proofs of the existence of universal constants for Feigenbaum period-doubling, developed by Sullivan, McMullen and Lyubich.

A *polynomial-like map* $f : U \to V$ is a proper holomorphic map between discs such that $\bar{U}$ is a compact subset of V. The *filled Julia set* $K(f)$ is defined to be

$$K(f) = \bigcap_{1}^{\infty} f^{-n}(V)$$

Two polynomial-like maps are said to be *hybrid equivalent* if there is a quasiconformal conjugacy ϕ between f and g, defined on a neighbourhood of their filled Julia sets, such that $\partial\phi/\partial\bar{z} = 0$ on $K(f)$.

Theorem (the "Straightening Theorem" of Douady and Hubbard):
Every polynomial-like map f is hybrid equivalent to a polynomial map g of the same degree. When $K(f)$ is connected, the polynomial g is unique up to affine conjugation.

Another key ingredient in renormalization theory is that of a *Yoccoz puzzle*, which provides a way to cut up a neighbourhhood of the filled Julia set of a polynomial, using external rays and equipotentials, so that the Poincaré return map of certain key pieces become polynomial-like mappings. McMullen's book "Complex Dynamics and Renormalization" is a good starting point to read about this technique.

4.7 Suggestions for further reading

The following list concentrates on topics covered in these notes and their further development. Some of the references are easier to find than others. The list is somewhat random, and far from complete.

Introductory survey articles on iterated rational maps:
P Blanchard, Complex analytic dynamics on the Riemann sphere, Bull AMS 11(1984) 85-141
M Lyubich, The dynamics of rational transforms: the topological picture, Uspekhi Mat Nauk 41 (1986) 35-95; Russian Math Surveys 41 (1986) 43-117

Textbooks on iterated rational maps (Fatou-Julia theory and selected topics from more recent work):
A Beardon, Iteration of rational functions, Springer Graduate Texts in Mathematics No.132, 1991
L Carleson and T Gamelin, Complex Dynamics, Springer Universitext 1993
J Milnor, Dynamics in one complex variable, third edition, Princeton University Press, January 2006

An advanced textbook on holomorphic dynamics, including rational maps, entire maps, Kleinian groups and iteration in $\mathbf{C}^n$:
S. Morosawa, Y. Nishimura, M.Taniguchi and T. Ueda, Holomorphic Dynamics, CUP, 2000

Textbooks on various aspects of hyperbolic geometry and Kleinian groups:

A Beardon, The geometry of discrete groups, Springer Graduate Text no. 91, 1983

B Maskit, Kleinian groups, Gpringer Grundlehren der mathematische Wissenschaft 287, 1987

J Ratcliffe, Foundations of hyperbolic manifolds, Springer Graduate Text no. 149, 1994

W Thurston, Three-dimensional geometry and topology, Princeton University Press, 1997

D Mumford, C Series and D Wright, Indra's pearls: the vision of Felix Klein, CUP, 2002

The work of Douady, Hubbard and their colleagues on the dynamics of polynomial maps, and polynomial-like mappings:

A Douady and J Hubbard, Itération des polynômes quadratiques complexes, Comptes Rendus Acad Sci Paris 294 (1982) 123-126

A Douady and J Hubbard, Etude dynamique des polynômes complexes (I et II), Publications Mathématiques d'Orsay, 1984/1985

A Douady and J Hubbard, On the dynamics of polynomial-like mappings, Ann. Ecole Norm. Sup. 18 (1985), 287-343

Sullivan's papers on quasiconformal deformation theory, including the proof of the 'No wandering Domains Theorem':

D Sullivan, Quasiconformal homeomorphisms and dynamics I: solution of the Fatou-Julia problem on wandering domains, Annals of Math 122 (1985) 401-418

D Sullivan, Quasiconformal homeomorphisms and dynamics II, Acta Math. 155 (1985), 24-260

D Sullivan and C T McMullen, Quasiconformal homeomorphisms and dynamics III, Adv. Math. 135 (1998), 351-395

The linearisation theorem of Siegel, Brjuno and Yoocoz:

J-C Yoccoz, Petits diviseurs en dimension 1, Astérisque 231, 1995

McMullen's work on renormalization and on relations between the theories of rational maps and Kleinian groups:

C T McMullen, Frontiers in complex dynamics, Bull AMS 31 (1994) 155-172

C T McMullen, Complex dynamics and renormalization, Annals of Math Studies No.135 (1994)

C T McMullen, The classification of conformal dynamical systems, *in* Current Developments in Mathematics 1995 (International Press) 323-360

C T McMullen, Renormalization and 3-manifolds which fiber over the circle, Princeton University Press 1996

Real and complex 1-dimensional dynamics (including Sullivan's proof of Feigenbaum universality):

W de Melo and S van Strien, One-dimensional dynamics, Springer 1993

Perhaps the most elegant and general statement and proof to date of Feigenbaum universality:

M Lyubich, *Feigenbaum-Coullet-Tresser universality and Milnor's hairiness conjecture*, Annals of Math. 149, 1999, 319-420

Internal addresses in the Mandelbrot set:

C Penrose, Quotients of the shift associated with dendrite Julia sets, preprint 1994

E Lau and D Schleicher, Internal addresses in the Mandelbrot set and irreducibility of polynomials, SUNY Stony Brook preprint 1994

The devil's staircase:

S Bullett and P Sentenac, Ordered orbits of the shift, square roots and the devil's staircase, Math Proc Cam Phil Soc 115 (1994) 451-481

K Keller, Symbolic dynamics for angle-doubling on the circle III: Sturmian sequences and the quadratic map, Ergod Th and Dyn Syst 14 (1994) 787-805

A collection of interesting papers on various topics concerning the Mandelbrot set:

Tan lei (editor), The Mandelbrot Set, Theme and Variations, LMS Lecture Notes 274, CUP 2000